UNIVERSITI TEKNOLOGI MARA

DEVELOPMENT OF PROCESS FRAMEWORK FOR ONTOLOGY CONSTRUCTION THROUGH ONTOLOGY MAPPING

SYERINA AZLIN MD NASIR

Thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Computer and Mathematical Sciences

January 2016

CONFIRMATION BY PANEL OF EXAMINERS

I certify that a panel of examiners has met on 14th May 2015 to conduct the final examination of Syerina Azlin Md Nasir on her Doctor of Philosophy thesis entitled "Development of Process Framework for Onology Construction through Ontology Mapping Approach" in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examniners recommends that the student be awarded the relevant degree. The panel of Examiners was as follows:

Daud Mohamad, PhD
Professor
Faculty of Computer & Mathematical Sciences
Universiti Teknologi MARA
(Chairman)

Christos Papatheodorou, PhD
Associate Professor
Department of Archives, Library Science & Museology
Ionaian University, Greece
(External Examiner)

Ahmad Zaki Abu Bakar, PhD
Professor
Faculty of Information and Communication Technology
Universiti Technikal Malaysia Melaka (UTEM)
(External Examiner)

Muthukaruppan Annamalai, PhD
Associate Professor
Faculty of Computer & Mathematical Sciences
Universiti Teknologi MARA
(Internal Examiner)

SITI HALIJJAH SHARIFF, PhD
Associate Professor
Dean
Institute of Graduates Studies
Universiti Teknologi MARA
Date: 27 January 2016

AUTHOR'S DECLARATIONS

I declare that the work in the thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the result of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any other degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Syerina Azlin Md Nasir
Student I.D. No.	:	2008252856
Programme	:	Doctor of Philosophy (Information Technology) CS990
Faculty	:	Computer & Mathematical Sciences
Thesis	:	Development of Process Framework for Ontology Construction through Ontology Mapping
Signature of Student	:	.....
Date	:	January 2016

ABSTRACT

Digitization of cultural heritage collections, both within and outside museums has led to the establishment of a new field of theory and practice of the digital curation of cultural information. Cultural heritage ontology construction a practice in museum documentation, faces the challenge of dealing with the abundance of information while struggling to maintain the authenticity and preservation of the cultural knowledge. Similar to other knowledge domains, the existence of various ontology of cultural heritage knowledge and their differences has caused predicament in accessing or retrieving information. To overcome this predicament, the standard ontology, in this case the CIDOC CRM is used to consolidate between two or more local ontology through the consolidation with a standard (global) ontology through a process known as ontology mapping. The effort needed to map two ontologies with some content similarity but different structures while minimizing knowledge loss is challenging. Knowledge loss is a situation where the concepts of local ontology are excluded from the result of mapping to the global ontology. Automated mapping through the use various existing algorithms offers an efficient solution but suffers in knowledge loss. A compromised method is needed to balance between the construction efficiency and the knowledge loss. This research seeks to establish a framework for constructing cultural heritage ontology through semi-automated ontology mapping. Using the traditional Malays textile (TMT) as the domain knowledge for this work and scoping on the Malaysian batik, this research is carried out in three phases. In the first phase, a manual mapping process between an existing TMT Knowledge Model (local) and CIDOC CRM (global) was carried to produce a Malaysian batik heritage ontology (MBHO). This is also achieved due to active participation of batik and ontology experts in verifying the MBHO. In the second phase, several automated mapping tools were tested. The finding shows that the mapping tools able to produce a Malaysian Batik ontology close to MBHO with knowledge loss as the resulting mapping were mostly incorrect. In the third phase, the refinement of the processes undergone in phases one and two were made to deduce the actual steps carried out and the rules that govern the actions of ontology construction through ontology mapping process. Comparison between both processes showed that the semi-automated mapping can improve the efficiency of the construction while reducing knowledge loss. The work in this phase led to the construction of a process framework of ontology construction through semi-automation. This framework is then verified by an expert in framework development and two experts in ontology constructions. A revised framework was produced and is name as the Framework for Ontology Construction through Semi-Automated Mapping (MapOn). This new framework contributes towards the knowledge and practice of digital curation in general and ontology construction in specific. In addition the MBHO is a new ontology for the Malaysian cultural heritage. This research also produces empirical evidence of the phenomena of knowledge loss through automated mapping which supports the inclination towards semi-automated mapping.

ACKNOWLEGEMENTS

In the Name of Allah, I am grateful that I have completed my thesis. I am indebted to the Ministry of Higher Education (MOHE) for the scholarship award. It has been a journey of invaluable experience and learning and discovering knowledge. Many people I would like to express my gratitude for helping me in the pursuit of this PhD research. I wish it were possible to thank them all.

I am most grateful to my supervisor, Prof Dr Nor Laila Md Noor for her supportive and inspiring guide throughout this PhD. Her guidance helped me in the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD study.

A very big thank you goes to Martin Doerr and Danuri Sakijan for their invaluable inputs; Sabine Massmann and Aliaksandr Autayeu for the supports and permission to use COMA++ and S-Match respectively in this study.

My warmest thanks to my father, Dato' Md Nasir Haji Musa, my mother, Azlina Haji Alias, my mother in law Hajjah Zaradah Abdul Rahman, for their constant prayers; my siblings, in laws and their families for their encouragement and understanding.

Most of all, my deepest appreciation is specially dedicated to my dear husband Ermy Azraai Ibrahim, my children, Nur Diana Ameera, Danial Afeef, Danish Aleef and Darwish Aqeef for their love and great patience at all times.